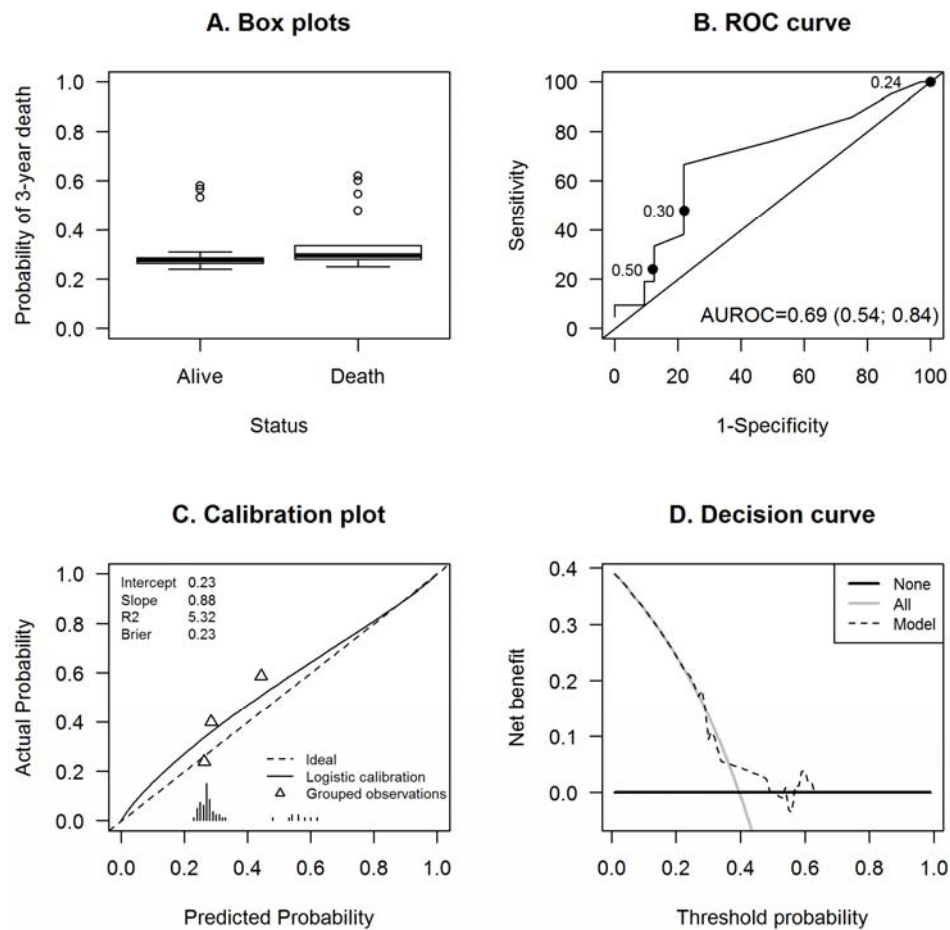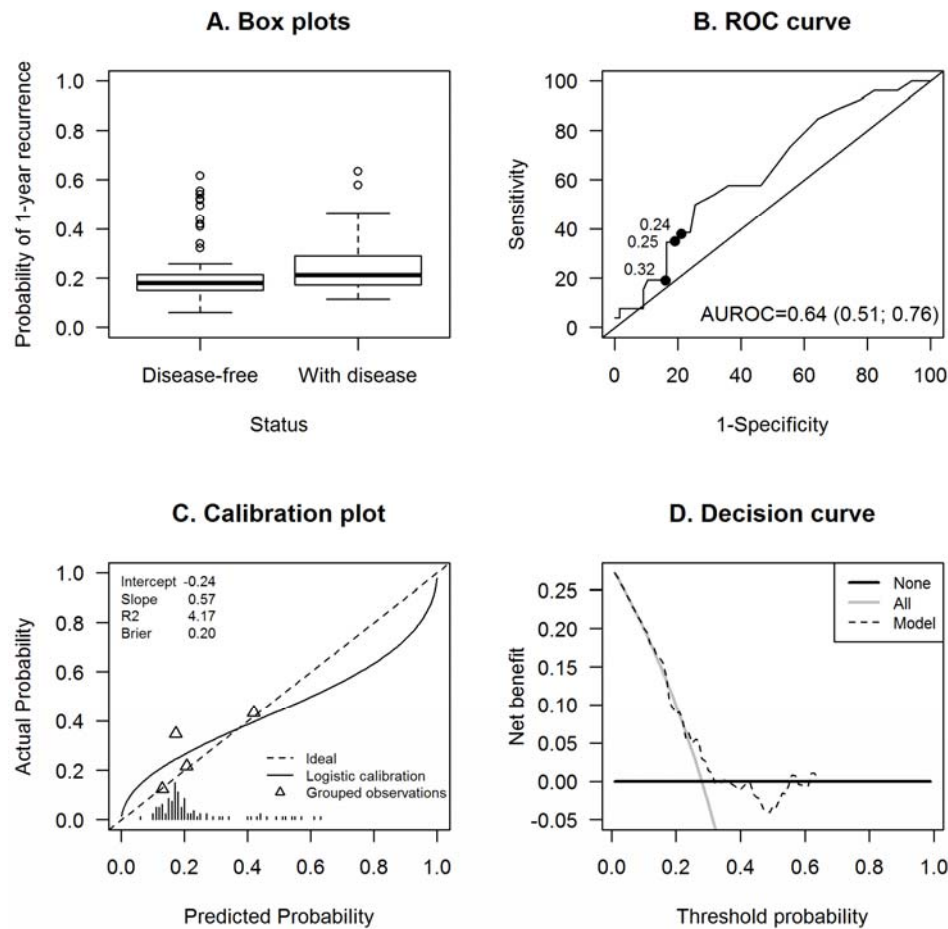**Supplementary 5.** Graphical approaches to visualize the predictive ability of the evaluated models
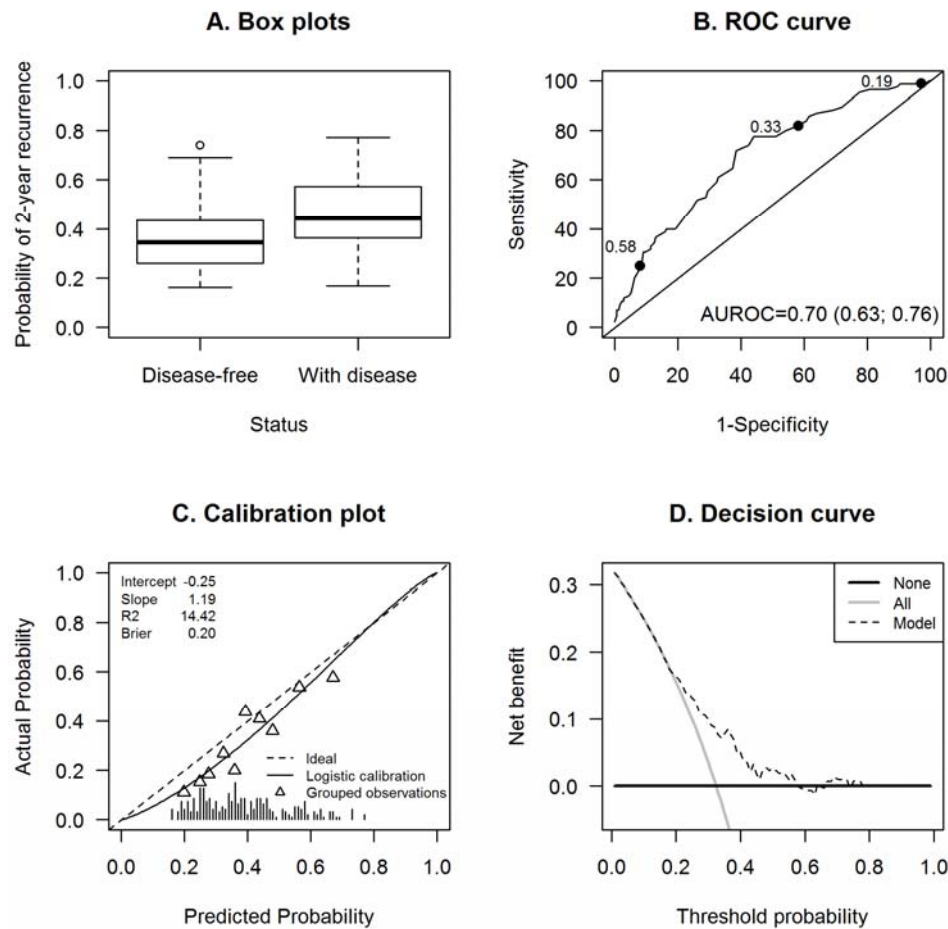

Supplementary Figure 1. Graphical plot for the OS model at 3 years from Kidd et al. [7]. (A) Box plots of predicted 3-years death risk for alive and death patients for OS model at 3 years from Kidd et al. (2012). (B) ROC curve with predicted probabilities closest to indifferent thresholds: 24% (NB all = NB threshold), 30% (NB all=NB none), and 50% (NB threshold=NB none). (C) Calibration plot: predicted outcome probabilities (on the x-axis) plotted against observed outcome frequencies (on the y-axis). Perfect predictions should be on the 45° line. (D) Decision curves for OS model at 3 years from Kidd et al. (2012) to identify the indifference thresholds.

NB, non-binary; OS, overall survival; ROC, receiver operating characteristic.
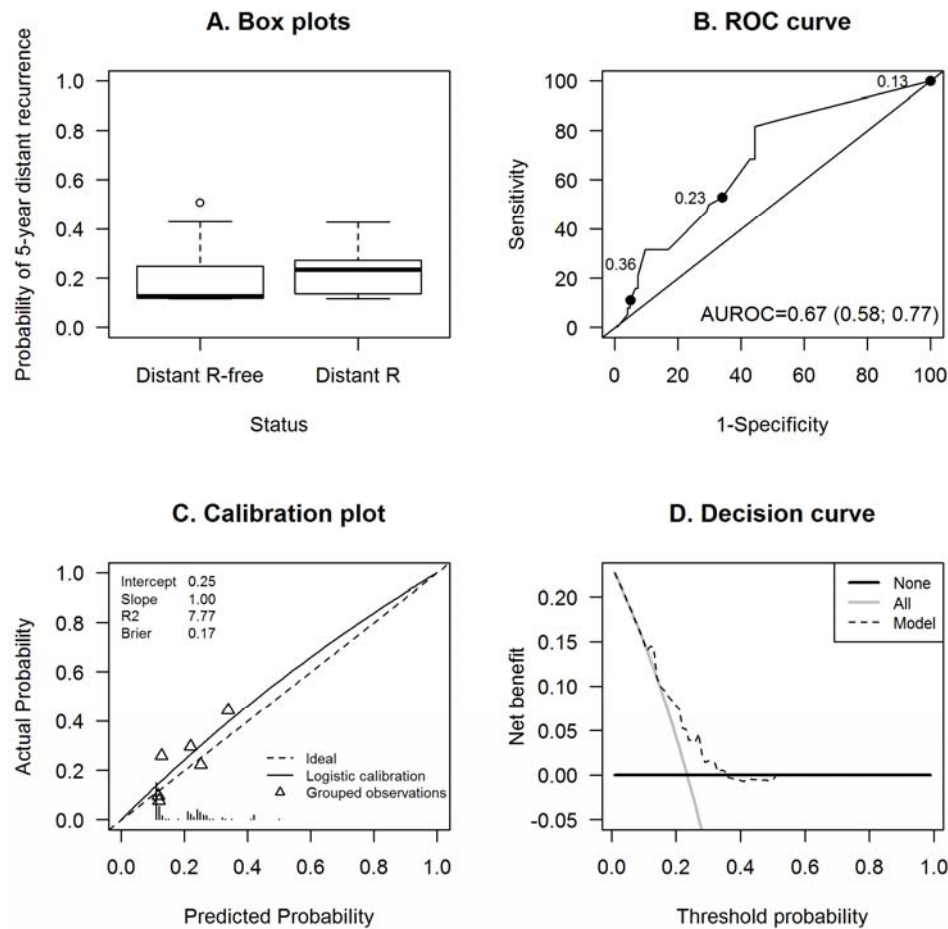
Supplementary Figure 2. Graphical plot for the DFS at 1 year from Kidd et al. [7]. (A) Box plots of predicted 1-years recurrence risk for without and with disease patients for DFS model at 1 years from Kidd et al. (2012). (B) ROC curve with predicted probabilities closest to indifferent thresholds: 24% (NB all=NB threshold), 25% (NB all=NB none), and 32% (NB threshold=NB none). (C) Calibration plot: predicted outcome probabilities (on the x-axis) plotted against observed outcome frequencies (on the y-axis). Perfect predictions should be on the 45° line. (D) Decision curves for DFS model at 1 years from Kidd et al. (2012) to identify the indifference thresholds.

DFS, disease-free survival; NB, non-binary; ROC, receiver operating characteristic.

Supplementary Figure 3. Graphical plot for the DFS at 2 years from Rose et al. [8]. (A) Box plots of predicted 2-years recurrence risk for without and with disease patients for DFS model at 2 years from Rose et al. (2013). (B) ROC curve with predicted probabilities closest to indifferent thresholds: 19% (NB all=NB threshold), 33% (NB all=NB none), and 58% (NB threshold=NB none). (C) Calibration plot: predicted outcome probabilities (on the x-axis) plotted against observed outcome frequencies (on the y-axis). Perfect predictions should be on the 45° line. (D) Decision curves for DFS model at 2 years from Rose et al. (2013) to identify the indifference thresholds.

DFS, disease-free survival; NB, non-binary; ROC, receiver operating characteristic.

Supplementary Figure 4. Graphical plot for the distant recurrence model at 5 years from Kang et al. [3]. (A) Box plots of predicted 5-years distant recurrence risk for without and with disease patients for distant recurrence model at 5 years from Kang et al. (2012). (B) ROC curve with predicted probabilities closest to indifferent thresholds: 13% (NB all=NB threshold), 23% (NB all=NB none), and 36% (NB threshold=NB none). (C) Calibration plot: predicted outcome probabilities (on the x-axis) plotted against observed outcome frequencies (on the y-axis). Perfect predictions should be on the 45° line. (D) Decision curves for distant recurrence model at 5 years from Kang et al. (2012) to identify the indifference thresholds.

NB, non-binary; ROC, receiver operating characteristic.